INFORMATION AND CONTROL 9, 393–413 (1966)

Optimal Control System Synthesis for Cost Functionals Involving Convex Single Valued Functions of the State and Control Variables*

D. ISAACS†

University of California, Irvine, California

AND

C. T. LEONDES†

University of California, Los Angeles, California

LIST OF SYMBOLS

$A(\cdot), B(\cdot), b(\cdot), D(\cdot)$	Matrices of coefficients
A^T, ψ^T	Transpose of a matrix or a vector
$x(\cdot, \cdot)$	State vector
$\mathbf{z}(\cdot, \cdot)$	Augmented state vector
x_0	Initial condition for the state vector
x_1	Final condition for the state vector
$\psi(\cdot, \cdot)$	Adjoint vector
$\boldsymbol{\psi}(\cdot, \cdot)$	Augmented adjoint vector
λ^*, λ, η	Initial condition for the adjoint vector
$u(\cdot, \cdot), u^*, v$	Control vectors
$X(\cdot)$	Solution to a matrix differential equation
$X^{-1}(\cdot)$	Inverse of the matrix $X(\cdot)$
$\varnothing$	Null matrix
I	Identity matrix
$\left.\begin{array}{l}\dfrac{\partial\psi}{\partial\lambda}, \dfrac{\partial x}{\partial\lambda}, \dfrac{\partial u}{\partial\psi} \\ \dfrac{\partial^2 H}{\partial x \partial\psi}, \dfrac{\partial^2 H}{\partial\psi\partial x}, \dfrac{\partial^2 H}{\partial x^2}, \dfrac{\partial^2 H}{\partial\psi^2}\end{array}\right\}$	Matrices of partial derivatives
$H(\cdot, \cdot, \cdot, \cdot)$	The max function
$\xi_{n+1}(\cdot, \cdot)$	A mapping function

* This research was supported under Grant AFOSR 699-65.

† The authors would like to acknowledge the reviewer and the excellent suggestions he made for the improvement of this paper.

τ	Fixed time of integration
t^l	The time at which one component of the control vector leaves or arrives at the boundary of U
U	The control region
E_n	n-dimensional Euclidean space
$\| x \|$	The norm of the vector x
$\bigcup$	The union of sets
$\Omega(\tau)$	The set of attainability
Γ	An open set

Neustadt (1961) has developed an algorithm for solving the time optimal problem where the system dynamics are linear in both the state and control variables, the control variables are bounded, and fixed or variable end point conditions are considered. He (Neustadt, 1963) later extended this work to include cost functionals which include the control variables. This approach is applicable to many control problems. Meditch and Neustadt (1964) used this method for a mid-course guidance scheme which minimized the total fuel expended during the guidance maneuver.

The purpose of this paper is to show how this technique can be extended to include at least quadratic terms in both the state and control variables in the cost functional. This result is not only of utility for control systems of the stated form, but is also of fundamental importance in two algorithmic approaches to optimal control systems synthesis problems of more general nature which are under investigation by the authors. In these techniques, the more general problem of controlling a system described by a set of first order nonlinear differential equations in an optimal manner according to a general cost function is treated by algorithms of "Sequential Optimization." In these algorithms, an initial estimate of the control vector as a function of time is sequentially improved using, among other techniques, the results developed in this paper.

I. PROBLEM STATEMENT

Given the dynamical equations:

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{1.1}$$

where:

$$
\begin{aligned}
&A(t) \text{ is an } n \times n \text{ matrix} \\
&B(t) \text{ is an } n \times m \text{ matrix} \\
&x(t) \text{ is an } n \times 1 \text{ vector} \\
&u(t) \text{ is an } m \times 1 \text{ vector} \\
&m \leqq n
\end{aligned}
\tag{1.2}
$$

with fixed boundary conditions:

$$
\begin{gathered}
x(0) = x_0, \qquad x(\tau) = x_1 \\
0 \leqq t \leqq \tau, \qquad \tau \text{ is a fixed value.}
\end{gathered}
\tag{1.3}
$$

We define a control region $U \subset E_m$. The set U is assumed compact. Vector valued control functions $u(t)$ are admissible, if each component is measurable and is in the range of U for $t \in [0, \tau]$.

We must find an admissible $u(t) \in U$ such that the cost functional

$$
x_{n+1}(\tau) = \int_0^\tau \{f(x, t) + h(u, t)\}\, dt \tag{1.4}
$$

is minimized. The constraints assumed in the integrand are:

(a) $f(x, t)$ and $h(u, t)$ are convex, single valued continuous functions of their respective arguments x and u for all $t \in [0, \tau]$, and $h(u, t)$ is strictly convex in u.

(b) The second partial derivatives of $f(x, t)$ and $h(u, t)$ exist and are continuous with respect to their arguments for all $t \in [0, \tau]$.
It is further assumed that $A(t)$ and $B(t)$ are continuous and bounded for all $t \in [0, \tau]$.

II. DERIVATION OF THE OPTIMAL CONTROL

Pontryagin (1962) has proven that if there is an admissible control $u(t)$, $t \in [0, \tau]$, which yields a solution to the optimal problem, then there exists a nonzero, continuous vector function $\boldsymbol{\psi}(t) = \{\psi_1, \cdots, \psi_{n+1}\}$ corresponding to $u(t)$ and $x(t)$ such that for all $t \in [0, \tau]$ the function $H[\boldsymbol{\psi}(t), x(t), u(t), t]$ attains its maximum at the point $u = u(t)$, i.e.:

$$
M[\boldsymbol{\psi}(t), x(t), t] = \underset{u \in U}{\operatorname{Max}}\, H[\boldsymbol{\psi}(t), x(t), u(t), t] \tag{2.1}
$$

where

$$
H[\boldsymbol{\psi}(t), x(t), u(t), t] = \psi^T(t)(A(t)x(t) + B(t)u(t))
$$

$$+ \psi_{n+1}(t)[f(x(t), t) + h(u, t)]$$
$$\psi = (\psi_1(t), \psi_2(t), \cdots \psi_n(t))^T$$
$$\boldsymbol{\psi} = (\psi(t), \psi_{n+1}(t)]^T.$$

The vector valued functions $x(t)$ and $\psi(t)$ must satisfy the differential equations given below.

$$\dot{x} = \frac{\partial H}{\partial \psi} = A(t)x + B(t)u \tag{2.2}$$

$$\dot{\psi} = -\frac{\partial H}{\partial x} = -\psi_{n+1}\frac{\partial f(x, t)}{\partial x} - [A(t)]^T\psi \tag{2.3}$$

$$\dot{\psi}_{n+1} = 0$$

$$\psi_{n+1} = \text{constant} \leqq 0.$$

We shall consider here the nondegenerate problem of Rozonoer; Thus we can normalize the augmented adjoint vector $\boldsymbol{\psi}(t)$ such that $\psi_{n+1} = -1$ (Rozonoer, 1959).

We shall assume that $x(\tau) = x'$ is not on the boundary of the set $\Omega(\tau)$. If this were not true, the optimal control would be totally bounded over the time interval $[0, \tau]$. Such problems are not considered and the computational method discussed here is not valid in this case. Furthermore, as the final desired state $x(\tau)$ approaches the boundary of $\Omega(\tau)$, the number of iterations required for convergence will increase rapidly until failure to converge occurs for $x(\tau)$ on the boundary of $\Omega(\tau)$.

Let $\lambda^* = (\lambda, -1) = (\lambda_1, \lambda_2, \cdots, \lambda_n, -1)$ be an arbitrary initial condition vector for the vector function $\boldsymbol{\psi}(t)$ in (2.3). We will define a real valued function $\xi_{n+1}(\lambda)$ such that $\xi_{n+1}(\lambda)$ attains its maximum at those values of λ for which the solution of (2.2) and (2.3) determines the optimal control. The geometrical significance of $\xi_{n+1}(\lambda)$ will be clarified below. (See, for instance, Fig. 1.) Let this value of λ be $\eta = (\eta_1, \eta_2, \cdots, \eta_n)$.

The general solution to (2.2) is given by

$$x(t) = X(t)\left\{x_0 + \int_0^t X^{-1}(s)B(s)u(s)\,ds\right\}. \tag{2.4}$$

$X(t)$ is the $n \times n$ matrix solution satisfying the following equations:

$$\dot{X}(t) = A(t)X(t)$$
$$X(0) = I \text{ (the identity matrix).}$$

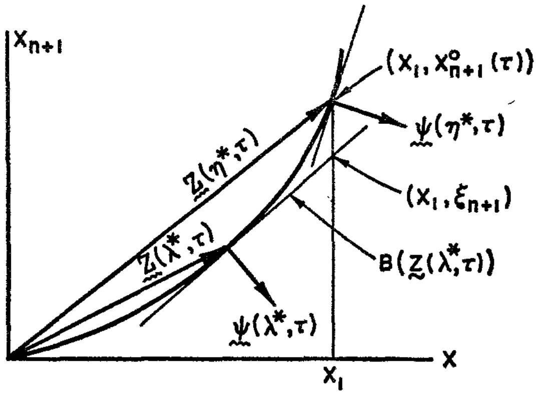

FIG. 1. The set of attainability

Let us define the set of attainability, $\Omega(t)$, as the set of all those end-points of the responses $(x(\tau), x_{n+1}(\tau))$ which are reachable, using an admissible control, from the initial condition $(x_0, 0)$.

$$\Omega(t) = \left\{ \begin{array}{l} X(t)\left[x_0 + \int_0^\tau X^{-1}(t)B(t)u(t)\,dt\right], \\ \qquad\qquad\qquad\qquad \int_0^\tau [f(x,t) + h(u,t)]\,dt; \\ \text{where: } u(t) \text{ is admissible.} \end{array} \right\} \tag{2.5}$$

The control $u(\lambda^*, t)$ is an extremal control if for some $\lambda^* = (\lambda, -1)$ it satisfies the maximum condition:

$$\begin{aligned} \psi(\lambda^*, t)\cdot B(t)u(\lambda^*, t) - h[u(\lambda^*, t), t] \\ = \operatorname*{Max}_{v\in U} \{\psi(\lambda^*, t)\cdot B(t)v - h(v, t)\} \end{aligned} \tag{2.6}$$

for all $t \in [0, t]$. Since $h(v, t)$ is strictly convex, $u(\lambda^*, t)$ is a unique maximum. $x(\lambda^*, t)$ and $\psi(\lambda^*, t)$ are defined to be the solutions of (2.2) and (2.3) with the initial conditions $x(\lambda^*, 0) = x_0$ and $\psi(\lambda^*, 0) = \lambda$.

In order that the problem not be vacuous, we shall assume that there are infinitely many extremal controls which will transfer the state vector from x_0 to the attainable set in the time τ. That extremal control which transfers the state vector from x_0 to x_1 and minimizes the cost functional $x_{n+1}(\tau)$ is called an optimal control.

Let us define the extremal solution vector $\mathbf{z}(\lambda^*, \tau)$ by

$$\mathbf{z}(\lambda^*, \tau) = (x(\lambda^*, \tau), x_{n+1}(\lambda^*, \tau)). \tag{2.7}$$

If the end point x_1 may be reached using an extremal control (Snow, 1964), then the final value of the optimal solution vector $\mathbf{z}(\eta^*, t)$ is given by:

$$\mathbf{z}(\eta^*, \tau) = (x_1, \overset{0}{x}_{n+1}(\tau)). \tag{2.8}$$

$\overset{0}{x}_{n+1}(\tau)$ is the minimum of $x_{n+1}(\tau)$. $\mathbf{z}(\eta^*, \tau)$ is the optimal solution vector which uses the optimal control function $u(\eta^*, t)$ on the interval $[0, \tau]$.

We now use Lee's Lemma (Lee, 1964) restated in our notation.

LEMMA. *Let $u(\lambda^*, t)$, $0 \leqq t \leqq \tau$ be an allowable extremal control with corresponding response $\mathbf{z}(\lambda^*, \tau)$ which starts at $(x_0, 0) = \mathbf{z}(\lambda^*, 0)$; for $\boldsymbol{\psi} = (\psi, \psi_{n+1})$, then:*

$$\boldsymbol{\psi}(\lambda^*, \tau) \cdot \mathbf{z}(\lambda^*, \tau) \geqq \boldsymbol{\psi}(\lambda^*, \tau) \cdot \boldsymbol{\omega}$$

for $\psi_{n+1} < 0$ and all $\omega \in \Omega(\tau)$.

If the set $\Omega(\tau)$ is closed, Lee (1964) has proven that the inequality of the Lemma establishes the fact that the set is convex. Furthermore, the outward pointing normal of the hyperplane which is tangent to the set is the final value of the adjoint vector $\boldsymbol{\psi}(\lambda^*, \tau)$ whose component $\psi_{n+1}(\lambda^*, \tau) < 0$.

In Lee's development, the inequality of the Lemma follows from Eq. (2.6) rewritten as follows:

$$\psi(\lambda^*, t) \cdot B(t) u(\lambda^*, t) - h[u(\lambda^*, t), t]$$
$$\geqq \psi(\lambda^*, t) \cdot B(t) u(t) - h(u(t), t)$$

for all $u(t) \in U$. Since $u(\lambda^*, t)$ is a unique maximum in (2.6), the $\geqq$ sign may be replaced by $>$ for all $u(t) \neq u(\lambda^*, t)$. Utilizing the convexity of $f(x, t)$ in Lee's development, we find that the equality will hold only if $\boldsymbol{\omega} = \mathbf{z}(\lambda^*, \tau)$. Thus we have strict convexity of the lower boundary of the set $\Omega(\tau)$.

The requirement that $h(u, t)$ be strictly convex and hence that Eq. (2.6) have a unique maximum certainly restricts the generality of the proposed method. Without this restriction the computational method using the gradient technique would not be applicable since strict convexity of $\Omega(\tau)$ would not be assured. Without the unique maximum condition it would not be possible to insure a unique optimal control. The existence of more than one optimal control would not necessarily prevent the existence of a unique optimal trajectory. Such cases are not considered in this study.

The Lemma has the geometrical representation shown in Fig. 1. In n-dimensional space, the curve would be a hypersurface and the tangent line would be a hyperplane.

Let $B(\mathbf{z}(\lambda^*, \tau))$ be the equation of this hyperplane which is tangent to the set at the point $\mathbf{z}(\lambda^*, \tau)$. Let $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_n, \xi_n, \xi_{n+1})$ be a vector whose end point lies on this plane at the point where the line $x(\tau) = x_1$ intersects this plane. We see that

$$\begin{aligned} \boldsymbol{\psi}(\lambda^*, \tau) \cdot \mathbf{z}(\lambda^*, \tau) &= \boldsymbol{\psi}(\lambda^*, \tau) \cdot \boldsymbol{\xi} \\ &= \boldsymbol{\psi}(\lambda^*, \tau) \cdot (x_1, \xi_{n+1}). \end{aligned} \tag{2.9}$$

From Lee's Lemma we have:

$$\boldsymbol{\psi}(\lambda^*, \tau) \cdot \mathbf{z}(\lambda^*, \tau) \geqq \boldsymbol{\psi}(\lambda^*, \tau) \cdot \mathbf{z}(\eta^*, \tau) \tag{2.10}$$

or:

$$\begin{aligned} \boldsymbol{\psi}(\lambda^*, \tau) \cdot (x_1, \xi_{n+1}) &\geqq \psi(\lambda^*, \tau) \cdot (x_1, \overset{0}{x}_{n+1}) \\ \psi(\lambda^*, \tau) \cdot x_1 - \xi_{n+1}(\tau) &\geqq \psi(\lambda^*, \tau) \cdot x_1 - \overset{0}{x}_{n+1}(\tau) \\ -\xi_{n+1}(\tau) &\geqq -\overset{0}{x}_{n+1}(\tau) \end{aligned} \tag{2.11}$$

or:

$$\xi_{n+1}(\tau) \leqq \overset{0}{x}_{n+1}(\tau).$$

If we solve (2.9) for $\xi_{n+1}(\tau)$

$$\begin{aligned} -\xi_{n+1} + \psi(\lambda^*, \tau) \cdot x_1 &= \boldsymbol{\psi}(\lambda^*, \tau) \cdot \mathbf{z}(\lambda^*, \tau) \\ \xi_{n+1} &= \psi(\lambda^*, \tau) \cdot x_1 - \boldsymbol{\psi}(\lambda^*, \tau) \cdot \mathbf{z}(\lambda^*, \tau). \end{aligned}$$

From (2.11) we see that in order to make $\xi_{n+1} = \overset{0}{x}_{n+1}$ we must maximize ξ_{n+1}. We see that ξ_{n+1} is actually a function of λ for a fixed τ. We indicate this by writing:

$$\xi_{n+1}(\lambda) = \psi(\lambda^*, \tau) \cdot x_1 - \boldsymbol{\psi}(\lambda^*, \tau) \cdot \mathbf{z}(\lambda^*, \tau). \tag{2.12}$$

Since the lower boundary of $\Omega(\tau)$ is strictly convex, the function $\xi_{n+1}(\lambda)$ will be a unique function of λ if $\boldsymbol{\psi}(\lambda^*, \tau)$ and $\mathbf{z}(\lambda^*, \tau)$ are unique functions of λ. The continuity of these functions with respect to their arguments will be discussed in Section V. The uniqueness of $\xi_{n+1}(\lambda)$ may be difficult to verify for particular cost functional integrands. In general, it can be said that $\xi_{n+1}(\lambda)$ will *not* be unique if the control determined from (2.6) is not unique.

III. THE METHOD OF STEEPEST ASCENT

We may use the method of steepest ascent to maximize $\xi_{n+1}(\lambda)$.] Since $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_n)$ is a vector in E_n, $\xi_{n+1}(\lambda)$ will trace out a path as we vary λ. It is necessary for steepest ascent that:

$$\frac{d\lambda_i}{ds} = k\,\frac{\partial \xi_{n+1}(\lambda)}{\partial \lambda_i} \qquad i = 1, 2, \cdots, n. \quad (3.1)$$

Since this is true for each component, we can see that we must have the following:

$$\frac{d\lambda}{ds} = k\nabla\xi_{n+1}(\lambda). \quad (3.2)$$

The constant k determines the rate at which we move along the gradient. In the discrete version, this is a step size constraint which must be determined either analytically or experimentally as we solve the iterative problem.

IV. EVALUATION OF THE GRADIENT

In order to evaluate the components of the gradient vector $\nabla\xi_{n+1}(\lambda)$, we must determine $\partial\xi_{n+1}(\lambda)/\partial\lambda_i$ from (2.12). Following Neustadt (1963), we can prove the following theorem.

THEOREM. *If $\boldsymbol{\psi}(\lambda^*, \tau) = (\psi(\lambda^*, \tau), -1)$ is the adjoint vector which is the outward normal to the tangent hyperplane at the point $\mathbf{z}(\lambda^*, \tau)$ of $\Omega(\tau)$, where $\lambda^* = (\lambda, -1)$ is the initial condition vector for the adjoint equations, and if $\partial\psi(\lambda^*, \tau)/\partial\lambda$ and $\mathbf{z}(\lambda^*, \tau)$ are both continuous at the point $\lambda^* = \lambda^0$ and $t = \tau$, then the function $\phi(\lambda) = \boldsymbol{\psi}(\lambda^*, \tau) \cdot \mathbf{z}(\lambda^*, \tau)$ has continuous partial derivatives at $\lambda^* = \lambda^0$ which are given by:*

$$\frac{\partial\phi(\lambda^0)}{\partial\lambda_i} = \frac{\partial\boldsymbol{\psi}(\lambda^0, \tau)}{\partial\lambda_i} \cdot \mathbf{z}(\lambda^0, \tau) \qquad i = 1, 2, \cdots, n. \quad (4.1)$$

Proof:

$$\phi(\lambda^*) - \phi(\lambda^0) = \boldsymbol{\psi}(\lambda^*) \cdot \mathbf{z}(\lambda^*) - \boldsymbol{\psi}(\lambda^0) \cdot \mathbf{z}(\lambda^0) \quad (4.2)$$

where we have dropped the τ for ease of notation. This can be written as

$$\phi(\lambda^*) - \phi(\lambda^0) = \boldsymbol{\psi}(\lambda^0) \cdot \mathbf{z}(\lambda^*) - \boldsymbol{\psi}(\lambda^0) \cdot \mathbf{z}(\lambda^0)$$
$$+ [\boldsymbol{\psi}(\lambda^*) - \boldsymbol{\psi}(\lambda^0)] \cdot \mathbf{z}(\lambda^0) + [\boldsymbol{\psi}(\lambda^*) - \boldsymbol{\psi}(\lambda^0)] \cdot [\mathbf{z}(\lambda^*) - \mathbf{z}(\lambda^0)].$$

Dividing both sides of this equation by $\| \lambda^* - \lambda^0 \|$:

$$\frac{\phi(\lambda^*) - \phi(\lambda^0)}{\| \lambda^* - \lambda^0 \|} = \frac{\psi(\lambda^0)\cdot \mathbf{z}(\lambda^*) - \psi(\lambda^0)\cdot \mathbf{z}(\lambda^0)}{\| \lambda^* - \lambda^0 \|}$$
$$+ \left[\frac{\psi(\lambda^*) - \psi(\lambda^0)}{\| \lambda^* - \lambda^0 \|}\right] \cdot z(\lambda^0) + \left[\frac{\psi(\lambda^*) - \psi(\lambda^0)}{\| \lambda^* - \lambda^0 \|}\right] \cdot [\mathbf{z}(\lambda^*) - \mathbf{z}(\lambda^0)]$$

Transposing and taking the absolute value:

$$\left| \frac{\phi(\lambda^*) - \phi(\lambda^0)}{\| \lambda^* - \lambda^0 \|} - \left[\frac{\psi(\lambda^*) - \psi(\lambda^0)}{\| \lambda^* - \lambda^0 \|}\right] \cdot \mathbf{z}(\lambda^0) \right| \leqq \left| \frac{\psi(\lambda^0)\cdot \mathbf{z}(\lambda^*) - \psi(\lambda^0)\cdot \mathbf{z}(\lambda^0)}{\| \lambda^* - \lambda^0 \|} \right| + \left| \left[\frac{\psi(\lambda^*) - \psi(\lambda^0)}{\| \lambda^* - \lambda^0 \|}\right] [\mathbf{z}(\lambda^*) - \mathbf{z}(\lambda^0)] \right| \tag{4.3}$$

From Lee's Lemma (1964), we have

$$\begin{aligned} \psi(\lambda^0)\cdot\ \mathbf{z}(\lambda^0) &\geqq \psi(\lambda^0)\cdot\ \mathbf{z}(\lambda^*) \\ \psi(\lambda^*)\cdot\ \mathbf{z}(\lambda^*) &\geqq \psi(\lambda^*)\cdot\ \mathbf{z}(\lambda^0) \end{aligned} \tag{4.4}$$

From (4.4) we see that

$$\begin{aligned} 0 &\geqq \psi(\lambda^0)\cdot\ \mathbf{z}(\lambda^*) - \psi(\lambda^0)\cdot\ \mathbf{z}(\lambda^0) \\ 0 &\geqq \psi(\lambda^*)\cdot\ \mathbf{z}(\lambda^0) - \psi(\lambda^*)\cdot\ \mathbf{z}(\lambda^*) \end{aligned} \tag{4.5}$$

From (4.5)

$$| \psi(\lambda^0)\cdot\ \mathbf{z}(\lambda^*) - \psi(\lambda^0)\cdot\ \mathbf{z}(\lambda^0) | \leqq | \psi(\lambda^0)\cdot\ \mathbf{z}(\lambda^*) - \psi(\lambda^0)\cdot\ \mathbf{z}(\lambda^0) + \psi(\lambda^*)\cdot\ \mathbf{z}(\lambda^0) - \psi(\lambda^*)\cdot\ \mathbf{z}(\lambda^*) |$$

Therefore:

$$\left| \frac{\psi(\lambda^0)\cdot \mathbf{z}(\lambda^*) - \psi(\lambda^0)\cdot \mathbf{z}(\lambda^0)}{\| \lambda^* - \lambda^0 \|} \right| \leqq \left| \left[\frac{\psi(\lambda^*) - \psi(\lambda^0)}{\| \lambda^* - \lambda^0 \|}\right] \cdot [\mathbf{z}(\lambda^*) - \mathbf{z}(\lambda^0)] \right|.$$

Substituting this relation into (4.3):

$$\left| \frac{\phi(\lambda^*) - \phi(\lambda^0)}{\| \lambda^* - \lambda^0 \|} - \left[\frac{\psi(\lambda^*) - \psi(\lambda^0)}{\| \lambda^* - \lambda^0 \|}\right] \cdot \mathbf{z}(\lambda^0) \right| \leqq 2 \left| \left[\frac{\psi(\lambda^*) - \psi(\lambda^0)}{\| \lambda^* - \lambda^0 \|}\right] \cdot [\mathbf{z}(\lambda^*) - \mathbf{z}(\lambda^0)] \right|. \tag{4.6}$$

Let us define $\lambda^* - \lambda^0 = (0 \cdots 0, \Delta_i\lambda, 0 \cdots 0)$, where $\Delta_i\lambda$ is the ith coordinate of $(\lambda^* - \lambda^0)$. Since $\| \lambda^* - \lambda^0 \| = | \Delta_i\lambda |$, (4.6) becomes

$$\left| \frac{\phi(\lambda^*) - \phi(\lambda^0)}{\Delta_i \lambda} - \left[\frac{\boldsymbol{\psi}(\lambda^*) - \boldsymbol{\psi}(\lambda^0)}{\Delta_i \lambda}\right] \cdot \mathbf{z}(\lambda^0) \right| \leqq 2 \left| \left[\frac{\boldsymbol{\psi}(\lambda^*) - \boldsymbol{\psi}(\lambda^0)}{\Delta_i \lambda}\right] \cdot [\mathbf{z}(\lambda^*) - \mathbf{z}(\lambda^0)] \right|.$$

Now taking the limit as $\Delta_i\lambda \to 0$, or $\lambda^* \to \lambda^0$, we see that:

$$\operatorname*{Lim}_{\Delta_i\lambda \to 0} \frac{\phi(\lambda^*) - \phi(\lambda^0)}{\Delta_i \lambda} = \frac{\partial\phi(\lambda^0)}{\partial\lambda_i}$$

$$\operatorname*{Lim}_{\Delta_i\lambda \to 0} \frac{\boldsymbol{\psi}(\lambda^*) - \boldsymbol{\psi}(\lambda^0)}{\Delta_i \lambda} = \frac{\partial\boldsymbol{\psi}(\lambda^0)}{\partial\lambda_i}$$

because of the continuity of $\mathbf{z}(\lambda^*)$ at $\lambda^* = \lambda^0$, $\mathbf{z}(\lambda^*) \to \mathbf{z}(\lambda^0)$; hence:

$$\frac{\partial\phi(\lambda^0)}{\partial\lambda_i} - \frac{\partial\boldsymbol{\psi}(\lambda^0)}{\partial\lambda_i} \cdot \mathbf{z}(\lambda^0) = 0.$$

Therefore we have:

$$\frac{\partial\phi(\lambda^0)}{\partial\lambda_i} = \frac{\partial\boldsymbol{\psi}(\lambda^0)}{\partial\lambda_i} \cdot \mathbf{z}(\lambda^0) = \frac{\partial\psi(\lambda^0)}{\partial\lambda_i} \cdot z(\lambda^0). \tag{4.7}$$

V. CONTINUITY WITH RESPECT TO THE INITIAL CONDITIONS

We found that it was necessary to maximize $\xi_{n+1}(\lambda)$ which was given by the following equation.

$$\xi_{n+1}(\lambda) = \psi(\lambda^*, \tau) \cdot x_1 - \boldsymbol{\psi}(\lambda^*, \tau) \cdot \mathbf{z}(\lambda^*, \tau). \tag{5.1}$$

From (3.2), we must evaluate $\nabla\xi_{n+1}(\lambda)$. Using (4.7), we find that

$$\frac{\partial\xi_{n+1}(\lambda)}{\partial\lambda_i} = \frac{\partial\psi(\lambda^*, \tau)}{\partial\lambda_i} \cdot (x_1 - z(\lambda^*, \tau)) \quad i = 1, 2, \cdots, n. \tag{5.2}$$

From the theorem in Section IV, the gradient is a continuous function if $\partial\psi(\lambda^*, \tau)/\partial\lambda$ and $\mathbf{z}(\lambda^*, \tau)$ are continuous functions of λ^* for a fixed τ. In order to demonstrate, therefore, that the gradient is a continuous function we will examine, in this section, the continuity of $\partial\psi(\lambda^*, \tau)/\partial\lambda$ and $z(\lambda^*, \tau)$. We will also digress, briefly, to examine the problem of the evaluation of the partial derivatives $\partial\psi(\lambda^*, \tau)/\partial\lambda$.

Considering first the continuity of $\partial\psi(\lambda^*, \tau)/\partial\lambda$, under the assumption that $h(u, t)$ is strictly convex in u, the control $u(\lambda^*, t)$, found from

(2.6), is a continuous function of $\psi(\lambda^*, t)$. Since $\psi(\lambda^*, t)$ is a continuous function of t, $u(\lambda^*, t)$ is continuous in t.

If $U = E_m$, the control $u(\lambda^*, t)$, determined from (2.6) can be replaced by a continuous function of $\psi(\lambda^*, t)$ in (2.2). We wish to show that the partial derivatives of x and ψ with respect to their initial conditions exist and are continuous on $t \in [0, \tau]$. An existence theorem may be found in Murray and Miller (1954), p. 86, under which the partial derivatives satisfy the required conditions.

THEOREM 2.

Hypothesis:

(i) *Let* $f_i(y_1 \cdots y_n, x), i = 1, \cdots, n$ *be* n *real valued functions of the* $n + 1$ *real variables* $f_1, y_2, \cdots y_n, x$ *defined and continuous on an open region* T *of* $(n + 1)$*-dimensional euclidean space.*

(ii) *Let* $\partial f_i/\partial y_i$ $i, j = 1, \cdots, n$ *exist and be jointly continuous in* $y_1, \cdots y_n, x$ *on* T.

(iii) *Let* $(y_{1,0}, \cdots, y_{n,0}, x_0)$ *be a point of* T.

(iv) *There exists a solution* $\{\theta_1(x), \cdots, \theta_n(x)\}$ *to the system of differential equations*

$$\frac{dy_i}{d\gamma} = f_i(y_1, \cdots, y_n, x), \qquad i = 1, \cdots, n$$

defined on $x_0 \leqq x \leqq c$ *with values* $\{\theta_1(x), \cdots \theta_n(x), x\}$ *in* T *and* $y_{j,0} = \theta_j(x_0)$.

Conclusion:

There exists a constant $b^+ > 0$ *and* n *functions*

$$y_i = \theta_i(x, y_1^*, \cdots y_n^*, x^*)$$

defined on a region M, $x_0 \leqq x \leqq c$, $|y_1^* - y_{i,0}| \leqq b^+$, $|x^* - x_0| \leqq b^+$ *such that*:

(i) $\partial\theta_i/\partial x = f_i(\theta_1, \cdots, \theta_n, x)$ *for* $x_0 \leqq x \leqq c$.

(ii) *The* $\theta_i(x, y_1^*, \cdots, y_n^*, x^*)$ *are jointly continuous in* $x_1 y_1^*, \cdots, y_n^*, x$ *on* M.

(iii) $y_i^* = \theta_i(x^*, y_1^*, \cdots, y_n^*, x^*)$.

(iv) *The* $\partial\theta_i/\partial y_j^*$ *exist and are continuous for*

$$x_0 \leqq x \leqq c, \qquad |y_i^* - y_{i,0}| < b^+, \qquad |x^* - x_0| \leqq b^+.$$

In the present application the right-hand sides of (2.2) and (2.3) correspond to $f_i(y_1, y_2, \cdots, y_n, x)$ and they satisfy the hypothesis of the theorem for $U = E_m$. Thus the partial derivatives exist and are continuous on $t \in [0, \tau]$.

Under these conditions, (5.2) represents a continuous function of λ. The problem remains as to the evaluation of the partial derivatives, $\partial\psi(\lambda^*, \tau)/\partial\lambda$. Scharmack (1963) has shown one way to evaluate similar partials with respect to the initial conditions. We will digress briefly to consider the problem of the evaluation of the partial derivatives, $\partial\psi(\lambda^*, \tau)/\partial\lambda$.

Let us consider formally the partial derivatives of the canonical equations (2.2) and (2.3) with respect to the initial condition vector λ.

$$\frac{\partial \dot{x}}{\partial \lambda} = \frac{\partial^2 H}{\partial x \partial \psi}\left(\frac{\partial x}{\partial \lambda}\right) + \frac{\partial^2 H}{\partial \psi^2}\left(\frac{\partial \psi}{\partial \lambda}\right),$$

$$\frac{\partial \dot{\psi}}{\partial \lambda} = -\frac{\partial^2 H}{\partial x^2}\left(\frac{\partial x}{\partial \lambda}\right) - \frac{\partial^2 H}{\partial \psi \partial x}\left(\frac{\partial \psi}{\partial \lambda}\right).$$

If we can interchange the partial and total derivatives, we have the following equations.

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial x}{\partial \lambda}\right) &= \frac{\partial^2 H}{\partial x \partial \psi}\left(\frac{\partial x}{\partial \lambda}\right) + \frac{\partial^2 H}{\partial \psi^2}\left(\frac{\partial \psi}{\partial \lambda}\right), \\ \frac{d}{dt}\left(\frac{\partial \psi}{\partial \lambda}\right) &= -\frac{\partial^2 H}{\partial x^2}\left(\frac{\partial x}{\partial \lambda}\right) - \frac{\partial^2 H}{\partial \psi \partial x}\left(\frac{\partial \psi}{\partial \lambda}\right). \end{aligned} \tag{5.3}$$

This interchange is valid if the following equalities hold (note that λ is not a function of time). Equations (5.3) are evaluated along the trajectories determined by the solution of (2.2) and (2.3) with their respective initial conditions

$$\frac{\partial^2 x}{\partial t \partial \lambda} = \frac{\partial^2 x}{\partial \lambda \partial t}$$

$$\frac{\partial^2 \psi}{\partial t \partial \lambda} = \frac{\partial^2 \psi}{\partial \lambda \partial t}.$$

If all derivatives are continuous in the domain of definition, the order of differentiation is immaterial (c.f. Kaplan (1953)). Thus (5.3) is valid. We can determine the initial conditions on (5.3) since $x(0) = x_0$, and $\psi(0) = \lambda$.

$$\left.\frac{\partial x}{\partial \lambda}\right|_{t=0} = \varnothing \quad \text{(null matrix)},$$

$$\left.\frac{\partial \psi}{\partial \lambda}\right|_{t=0} = I \quad \text{(identity matrix)}.$$

Note that $\partial x/\partial\lambda$ and $\partial\psi/\partial\lambda$ are $n \times n$ matrices. Let us define

$$\frac{\partial \psi}{\partial \lambda} = \begin{Bmatrix} \frac{\partial \psi_1}{\partial \lambda_1}, \frac{\partial \psi_1}{\partial \lambda_2}, \cdots, \frac{\partial \psi_1}{\partial \lambda_n} \\ \frac{\partial \psi_n}{\partial \lambda_1}, \frac{\partial \psi_n}{\partial \lambda_2}, \cdots, \frac{\partial \psi_n}{\partial \lambda_n} \end{Bmatrix} = Y^T. \tag{5.4}$$

Now (5.2) can be written in component notation.

$$\begin{aligned} \frac{\partial \xi_{n+1}(\lambda)}{\partial \lambda_i} &= \sum_{j=1}^{n} \frac{\partial \psi_j}{\partial \lambda_i} (x_1{}^j - z^j(\lambda^*, \tau)), \\ &= \sum_{j=1}^{n} Y_{ij}(x_1{}^j - z^j(\lambda^*, \tau)). \end{aligned} \tag{5.5}$$

$Y_{ij} = \partial\psi_j(\lambda^*, \tau)/\partial\lambda_i$ is one component of the solution of (5.3) and $z^j(\lambda^*, \tau)$ is the jth component of $\mathbf{z}(\lambda^*, \tau)$. In general we have the gradient determined.

$$\nabla\xi_{n+1}(\lambda) = Y(x_1 - z(\lambda^*, \tau)). \tag{5.6}$$

Using (3.2) and (5.6), we have our iteration defined by the following.

$$\frac{d\lambda}{ds} = kY(x_1 - z(\lambda^*, \tau)).$$

In the gradient technique proposed, we will use the discrete version.

$$\lambda^{(i+1)} = \lambda^{(i)} + kY(x_1 - z(\lambda^{(i)}, \tau)). \tag{5.7}$$

With any initial estimate $\lambda^{(i)}$, we can get an improved estimate since $z(\lambda^{(i)}, \tau)$ is just the result of the integration of our original set of equations.

The gradient technique is applicable if "k" is a sufficiently small constant and if $\xi_{n+1}(\lambda)$ has a unique extremum. If the extremum were not unique this would require that initial conditions λ be chosen such that $\xi_{n+1}(\lambda)$ be in a sufficiently close neighborhood of $x_{n+1}^0(\tau)$ for the gradient technique to be applied. Techniques for determining the optimal step size $k\nabla\xi_{n+1}(\lambda)$ are well known (Saaty and Bram, 1964). An acceleration method for convergence has been developed by Powell (1962) and applied successfully by Paiewonsky (1963).

If $U \subset E_m$, the control $u(\lambda^*, t)$ or some of its components can be bounded over some subintervals of $[0, \tau]$. We can break up the interval $[0, \tau]$ into subintervals I_i.

$$I_i = \{t \in [0, \tau]; 0 \leqq t^{i-1} \leqq t \leqq t^i \leqq \tau, i = 1, \cdots, N\}. \quad (5.8)$$

We have assumed a finite number of these intervals. Then

$$\bigcup_{i=1}^{N} I_i = [0, \tau].$$

The subintervals are taken such that t^i represents the time at which one or more components either reach the boundary from the interior of U or leave the boundary for the interior of U. Considering the first such interval I_1, we see that Theorem 2 (Murray and Miller, 1954) is satisfied for $t \in I_1$. We find that $x(\lambda^*, t)$ and $\psi(\lambda^*, t)$ have continuous partial derivatives with respect to the initial condition vector λ. Since $u(\lambda^*, t)$ is a continuous function of $\psi(\lambda^*, t)$, then the point $t = t^1$ must also be a continuous function of λ^*.

Let us indicate this by defining

$$x^1 = x(\lambda^*, t^1(\lambda^*)),$$
$$\psi^1 = \psi(\lambda^*, t^1(\lambda^*)).$$

Differentiation with respect to the initial conditions yields the equations which must be satisfied as initial conditions on the next interval.

$$\frac{\partial x_j^1}{\partial \lambda_j^1} = \frac{\partial x_i(\lambda^*, t)}{\partial \lambda_j} + \dot{x}_1(\lambda^*, t) \frac{\partial t^1(\lambda^*)}{\partial \lambda_j},$$
$$i, j = 1, 2, \cdots, n. \quad (5.9)$$
$$\frac{\partial \psi_i^1}{\partial \lambda_j^1} = \frac{\partial \psi_i(\lambda^*, t)}{\partial \lambda_j} + \dot{\psi}_i(\lambda^*, t) \frac{\partial t^1(\lambda^*)}{\partial \lambda_j},$$

The right-hand sides of (5.9) are to be evaluated at $t = t^1$. Continuing in this fashion for t^2, $t^3 \cdots$, t^N we can generate $\partial x(\lambda^*, \tau)/\partial x$ and $\partial \psi(\lambda^*, \tau)/\partial \lambda$ which are continuous functions of λ but only piecewise continuous in t. We notice from these equations that if $\partial t^1(\lambda^*)/\partial \lambda_j = 0$, then (5.9) reduce immediately to the set of continuous solutions of (5.3).

Considering the right-hand side of (5.9) again, the terms $\partial x_i(\lambda^*, t)/\partial\lambda_j$ and $\partial\psi_i(\lambda^*, t)/\partial\lambda_j$ are determined at times $t^1, t^2, \cdots, t^N$ from their value at the end of the previous interval as determined by (5.3). The values for the terms $\dot{x}_i(\lambda^*, t)$ and $\dot{\psi}_i(\lambda^*, t)$ are determined at the times $t^1, t^2, \cdots, t^N$ from (2.2) and (2.3). It remains to determine $\partial t^i(\lambda^*)/\partial\lambda_j$. The method for doing this may be found in Scharmack (1963) and is given explicitly by Eqs. (6.20), (6.27), and (8.22) of his paper. Briefly, the technique in Scharmack is, of course, based on the constraint equations on the control vector, which in vector form we may write as $G(x, u) \geqq 0$, where $G(x, u)$ is a "q" vector where "q" is the number of constraints involved, 0 is also a "q" vector, and x and u are the state and control vectors, respectively. If we assume that at time t^1 the constraint involved is the first element G^1 of the constraint vector G and if we replace the u vector in the G vector by x and ψ through the use of Equation (2.6), then t^1 is given from

$$G^1(x(t^1, \lambda^*), \psi(t^1, \lambda^*)) = 0.$$

If partial differentiation with respect to t^1 is taken on this equation then from implicit function theory the result may be solved to obtain $\partial t^1(\lambda^*)/\partial\lambda_j$ as shown by Eq. (8.22) in Scharmack (1963). When coming off a boundary the same general technique is used, and in this manner we may obtain $\partial t^i(\lambda^*)/\partial\lambda_j$. (See Scharmack's paper for further details and examples.)

Returning again to the question of the continuity of the gradient we consider now the demonstration of the continuity of $z(\lambda^*, \tau)$ with respect to λ^* for a fixed τ.

Over the entire interval $[0, \tau]$, $\psi(\lambda^*, t)$ satisfies the conditions of Theorem 2 (Murray and Miller, 1954). Thus $\psi(\lambda^*, t)$ is a continuous function of both λ and t. The requirement that $\dot{x}$ and $\dot{\psi}$ be defined on an open region is met if we define Γ as shown.

$$\Gamma = \{(x, \psi) \in E_n \times E_n ; x(\lambda, \tau) \in \dot{\overline{\Omega(\tau)}}, \forall\, t \in [0, \tau]\}.$$

$\dot{\overline{\Omega(\tau)}}$ is the interior of $\Omega(\tau)$.

It is obvious that $x(\lambda^*, t)$ is continuous in t. Lee's Lemma (1964) has shown that for $\mathbf{z}(\lambda^*, \tau) \in \Omega(\tau)$, the following inequality holds.

$$\boldsymbol{\psi}(\lambda^*, \tau) \cdot \mathbf{z}(\lambda^*, \tau) \geqq \boldsymbol{\psi}(\lambda^*, \tau) \cdot \boldsymbol{\omega}$$

for all $\boldsymbol{\omega} \in \Omega(\tau)$.

Since the control set $U \subset E_m$ was assumed compact and the time interval, $[0, \tau]$, is finite, the set $\Omega(\tau)$ is bounded. This follows directly from our assumptions on $f(x, t)$, $h(u, t)$, $A(t)$, and $B(t)$. (See Aoki (1964).)

Let $\{\lambda_j\} \in E_{n+1}$ be a sequence of vectors such that

$$\lim_{j\to\infty} \lambda_j = \lambda^*.$$

Then the set $\{\mathbf{z}(\lambda_j, \tau)\}$ has at least one cluster point. Since every bounded sequence in E_{n+1} has a convergent subsequence, we define this subsequence as follows:

$$\lim_{k\to\infty} \lambda_{jk} = \lambda^*,$$

$$\lim_{k\to\infty} \mathbf{z}(\lambda_{jk}, \tau) = \beta^*.$$

Let us assume there is a second such subsequence

$$\lim_{l\to\infty} \lambda_{jl} = \lambda^*,$$

$$\lim_{l\to\infty} \mathbf{z}(\lambda_{jl}, \tau) = \beta^{**}.$$

Using Lee's Lemma we have

$$\psi(\lambda_{jk}, \tau)\cdot \mathbf{z}(\lambda_{jk}, \tau) \geqq \psi(\lambda_{jk}, \tau)\cdot \mathbf{z}(\lambda_{jl}, \tau),$$

$$\psi(\lambda_{jl}, \tau)\cdot \mathbf{z}(\lambda_{jl}, \tau) \geqq \psi(\lambda_{jl}, \tau)\cdot \mathbf{z}(\lambda_{jk}, \tau).$$

Since $\psi(\lambda, \tau)$ is continuous in λ, we can take the limit of the first equation with respect to k and the second with respect to l.

$$\psi(\lambda^*, \tau)\cdot\beta^* \geqq \psi(\lambda^*, \tau)\cdot \mathbf{z}(\lambda_{jl}, \tau),$$

$$\psi(\lambda^*, \tau)\cdot\beta^{**} \geqq \psi(\lambda^*, \tau)\cdot \mathbf{z}(\lambda_{jk}, \tau).$$

Again taking limits with respect to l and k,

$$\psi(\lambda^*, \tau)\cdot\beta^* \geqq \psi(\lambda^*, \tau)\cdot\beta^{**},$$

$$\psi(\lambda^*, \tau)\cdot\beta^{**} \geqq \psi(\lambda^*, \tau)\cdot\beta^{*}.$$

This implies $\beta^* = \beta^{**} = \beta$, hence we have

$$\lim_{j\to\infty} \lambda_j = \lambda^*,$$

$$\lim_{j\to\infty} \mathbf{z}(\lambda_j, \tau) = \beta.$$

Again using the lemma,

$$\psi(\lambda_j, \tau) \cdot \mathbf{z}(\lambda_j, \tau) \geqq \psi(\lambda_j, \tau) \cdot \mathbf{z}(\lambda^*, \tau).$$

Consider the limit as $j \to \infty$

$$\psi(\lambda^*, \tau) \cdot \beta \geqq \psi(\lambda^*, \tau) \cdot \mathbf{z}(\lambda^*, \tau).$$

Since $\Omega(\tau)$ is strictly convex, this is possible, if and only if,

$$\beta = \mathbf{z}(\lambda^*, \tau).$$

This implies that $\mathbf{z}(\lambda^*, \tau)$ is continuous at $\lambda = \lambda^*$, since $\mathbf{z}(\lambda^*, \tau)$ is bounded for any sequence of vectors $\{\lambda_j\}$ which converges to λ^*.

VI. EVALUATION OF THE PARTIAL DERIVATIVES

The use of this method requires the solution of the set of (5.3). They can be integrated along with the dynamical equations which yield $z(\lambda^*, \tau)$. If we do this, the matrix of coefficients $\partial^2 H/\partial x^2$ and $\partial^2 H/\partial \psi^2$ is already calculated. At the final time $t = \tau$, the solution of this set of equations yields the desired $\partial\psi/\partial\lambda$ for our use in (5.6).

For this particular problem, the coefficients of (5.3) are:

$$\begin{aligned}
\frac{\partial^2 H}{\partial x\, \partial \psi} &= A(t), \\
\frac{\partial^2 H}{\partial \psi\, \partial x} &= [A(t)]^T, \\
\frac{\partial^2 H}{\partial x^2} &= -\frac{\partial^2 f(x, t)}{\partial x^2}.
\end{aligned} \tag{6.1}$$

All these terms are known and can be readily evaluated. The term $\partial^2 H/\partial\psi^2$ can be evaluated from the following.

$$\begin{aligned}
\dot{x} &= A(t)\, x + B(t)\, u, \\
\dot{x} &= \frac{\partial H}{\partial \psi}.
\end{aligned}$$

Let us take partial derivatives with respect to the initial condition vector λ as in (5.3).

$$\begin{aligned}
\frac{d}{dt}\left(\frac{\partial x}{\partial \lambda}\right) &= A(t)\left(\frac{\partial x}{\partial \lambda}\right) + B(t)\left(\frac{\partial u}{\partial \lambda}\right), \\
\frac{d}{\partial t}\left(\frac{\partial x}{\partial \lambda}\right) &= \frac{\partial^2 H}{\partial x\, \partial \psi}\left(\frac{\partial x}{\partial \lambda}\right) + \frac{\partial^2 H}{\partial \psi^2}\left(\frac{\partial \psi}{\partial \lambda}\right).
\end{aligned}$$

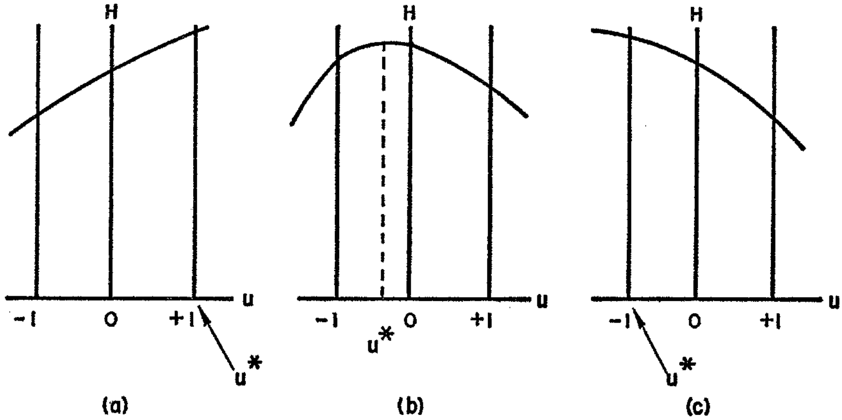

FIG. 2. Possible variation of the max function with control for fixed time.

Since

$$\frac{\partial u}{\partial \lambda} = \frac{\partial u}{\partial \psi} \frac{\partial \psi}{\partial \lambda}$$

we must have

$$\frac{\partial^2 H}{\partial \psi^2} = B(t) \frac{\partial u}{\partial \psi}. \tag{6.2}$$

$u(t)$ is determined from (2.6) and hence is a function of ψ. H is a concave function of u for $h(u, t)$ is a convex function of u for all $t \in [0, \tau]$. Since the set U is a compact set, the maximizing $u = u^*$ will be found either on the boundary of the set or in its interior for any fixed t. In two-dimensional form, these possibilities would be as shown in Fig. 2.

In order to illustrate more clearly the procedure, a particular form of $h(u, t)$ is used which is amenable to analysis.

$$h(u, t) = b(t)u + u^T D(t)u.$$

Both $b(t)$ and $D(t)$ are assumed $m \times m$ diagonal matrices which are continuous for $t \in [0, \tau]$. The different cases are shown in Fig. 2.

CASE (b)

We can determine the maximizing $u = u^*$ by taking partial derivatives of H with respect to u and equating to zero.

$$H = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_j\, \psi_i + \sum_{i=1}^{n} \sum_{k=1}^{m} b_{ik}\, u_k\, \psi_i - f(x,t) - \sum_{k=1}^{m} b_k\, u_k - \sum_{k=1}^{m} D_k u_k^{\,2},$$

$$\frac{\partial H}{\partial u_k} = \sum_{i=1}^{n} b_{ik}\, \psi_i - b_k - 2D_k\, u_k = 0.$$

The maximizing u_k^* is given by

$$u_k^* = \sum_{i=1}^{n} \frac{b_{ik}}{2D_k} \psi_i - \frac{b_k}{2D_k}. \tag{6.3}$$

From (6.3), we can also determine the required partial derivatives.

$$\frac{\partial u_k^*}{\partial \psi_i} = \frac{b_{ik}}{2D_k}. \tag{6.4}$$

From (6.4), we can develop the following.

$$\frac{\partial u}{\partial \psi} = \begin{Bmatrix} \dfrac{\partial u_1}{\partial \psi_1} & \cdots & \dfrac{\partial u_1}{\partial \psi_n} \\ \dfrac{\partial u_m}{\partial \psi_1} & \cdots & \dfrac{\partial u_m}{\partial \psi_n} \end{Bmatrix} = \frac{1}{2} \begin{Bmatrix} \dfrac{b_{11}}{D_1} & \cdots & \dfrac{b_{n1}}{D_1} \\ \dfrac{b_{1m}}{D_m} & \cdots & \dfrac{b_{nm}}{D_m} \end{Bmatrix}$$

or:

$$\frac{\partial u}{\partial \psi} = \frac{1}{2} \begin{Bmatrix} \dfrac{b_{11}}{D_1} & \cdots & \dfrac{b_{1m}}{D_m} \\ \vdots & & \\ \dfrac{b_{n1}}{D_1} & \cdots & \dfrac{b_{nm}}{D_1} \end{Bmatrix}^T = \frac{1}{2} \left[\begin{Bmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \\ b_{n1} & \cdots & b_{nm} \end{Bmatrix} \begin{Bmatrix} \dfrac{1}{D_1} & 0 & \cdots & 0 \\ 0 & \dfrac{1}{D_2} & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & 0 & \dfrac{1}{D_m} \end{Bmatrix} \right]^T$$

$$= \frac{1}{2} (BD^{-1})^T = \frac{1}{2} (D^{-1})^T B^T$$

$$\frac{\partial u}{\partial \psi} = \frac{1}{2} D^{-1} B^T. \tag{6.5}$$

From (6.2), we thus have:

$$\frac{\partial^2 H}{\partial \psi^2} = \frac{1}{2} BD^{-1}B^T. \tag{6.6}$$

Using (6.1) and (6.6) in (5.3), we have:

$$\frac{d}{dt}\left(\frac{\partial x}{\partial \lambda}\right) = A(t)\left(\frac{\partial x}{\partial \lambda}\right) + \frac{1}{2} BD^{-1}B^T \left(\frac{\partial \psi}{\partial \lambda}\right)$$
$$\frac{d}{dt}\left(\frac{\partial \psi}{\partial \lambda}\right) = \frac{\partial^2 f(x)}{\partial x^2}\left(\frac{\partial x}{\partial \lambda}\right) - A^T(t)\left(\frac{\partial \psi}{\partial \lambda}\right). \tag{6.7}$$

Thus in any time interval $[a, b]$, where $0 \leqq a \leqq t \leqq b \leqq \tau$, if u^* falls within the interior of U, we solve (6.7) by integration.

Case (a) or (c)

If in a finite time interval $I_i \in [0, \tau]$, one or more components of u remain on the boundary of U, then the partial derivatives of these components with respect to ψ are identically zero. The rest of the components u_k^* are determined from (6.3) and u_i^* are set equal to the boundary values for $i \neq k$.

VII. THE CLOSURE OF THE SET $\Omega(\tau)$

The existence of the optimal control is dependent upon the closure of the set of attainability, if there is at least one admissible control which steers the system from the given initial condition to the desired end condition. Lee (1964) has pointed out that the set $\Omega(\tau)$ is known to be closed when the cost functional integrand is of positive definite quadratic form in the control and state variables. This can also be extended to include those intergrands which are strictly convex in the control variables.

Consider a sequence $\{\lambda_j\} \in E_{n+1} \ni$

$$\lim_{j\to\infty} \lambda_j = \lambda^*. \tag{7.1}$$

Then any sequence $\{z(\lambda_j, \tau)\} \in \Omega(\tau) < E_{n+1}$ which has a cluster point $\underline{\beta}$ must also have a subsequence $\{z(\lambda_{jk}, \tau)\} \ni$

$$\lim_{k\to\infty} \mathbf{z}(\lambda_{jk}, \tau) = \underline{\beta} \tag{7.2}$$

and

$$\lim_{k\to\infty} \lambda_{jk} = \lambda^*. \tag{7.3}$$

But $\mathbf{z}(\lambda_{jk}, \tau)$ is bounded for all j and k with a fixed $\tau < \infty$. This implies that $\underline{\beta} \in \Omega(\tau)$. Since the sequence $\{\lambda_j\}$ was arbitrary, all cluster points such as $\underline{\beta}$ are contained in $\Omega(\tau)$. Hence $\Omega(\tau)$ is closed.

Received: October 28, 1965
Revised: November 23, 1965

REFERENCES

AOKI, M., (1964), On optimal and suboptimal policies in control systems. *Advanc. Control Systems* **1**, 43–51.

DIEUDONNÉ, J., (1960), "Foundations of Modern Analysis." Academic Press, New York.

KAPLAN, W., (1953), "Advanced Calculus." Addison-Wesley, Reading, Mass.

LEE, E. B., (1964), A sufficient condition in the theory of optimal control. *SIAM J. Control* **1**, No. 3.

MEDITCH, J. S., AND NEUSTADT, L. W., (1964), An application of optimal control to midcourse guidance. *Proc. IFAC Conf. Basle, Switzerland*, August 1964.

MURRAY, F. J., AND MILLER, K. S., (1954), "Existence Theorems for Ordinary Differential Equations." New York Univ. Press.

NEUSTADT, L. W., (1961), Time optimal control systems with position and integral limits. *J. Math. Anal. Appl.* **3**.

NEUSTADT, L. W., (1963), On synthesizing optimal controls. *Proc. IFAC Conf. Basle, Switzerland*, August 1963.

PAIEWONSKY, B. H., (1963), The synthesis of optimal controls for a class of rocket steering problems. *AIAA Summer Meeting*, June 1963.

PONTRYAGIN, L. S., (1962), "The Mathematical Theory of Optimal Processes." Interscience, New York.

POWELL, M. J. D., (1962), An iterative method for finding stationary values of a function of several variables. *Comput. J.* **5**, 147–151.

ROZONOER, L. I., (1959), L. S. Pontryagin maximum principle in the theory of optimum systems. I. *Automation Remote Control* **20**, No. 10.

SAATY, T. L., AND BRAM, J., (1964), "Nonlinear Mathematics." McGraw-Hill, New York.

SCHARMACK, D. K., (1963), "A Second Order Method for the Numerical Solution of the Control Optimization Problems," U-RD 6304, (Minneapolis-Honeywell, September 16, 1963). Also to appear in *Advanc. Control Systems* **4** (1966).

SNOW, D. R., (1964), "Reachable Regions and Optimal Controls," Ph.D. Dissertation, Department of Mathematics, Stanford University, October 1964.